Second Generation Bandwidth Engine® IC
Breaks 4.5 Billion Accesses/sec

Michael J. Miller
VP, Technology Innovation & System Applications
The Vision: Fast, Intelligent Access Architecture

Packet Processor

ingress

Multi-threaded Multi-Cores allow for high processing throughput

Multi-bank Multi-partitions allow for high access availability

Allows Extended Carrier Class & In package Repair

egress

Multi-linked allow for concurrent transport operations

Fixed Macro ALUs for functional Acceleration minimizes intra-chip traffic

Atomic Memory Operations

Bandwidth Engine – Second Generation

ALU

ALU

Bit Safe™ Technology

Serial Link

Serial Link

Serial Link

Serial Link

Copyright ©MoSys, Inc. 2013. All rights reserved.
Bandwidth Engine 2 Architecture & Family Sampling Now

- **Parallel Array Architecture**  ... Performance up to:
  - 16 outstanding transactions
  - 4.5G Accesses per second
  - 192 Gbps full duplex throughput
  - ~12ns deterministic read latency
  - 2.7ns Random cycle time (tRC)

- **GigaChip Interface**  ... 90% Efficient Transport Protocol
  - Up to sixteen low latency SerDes lanes (8G to 15G)

- **High Reliability**  ...70X better SER than 6T-SRAM
  - Full ECC support; 72bit array and macro datapath operations
  - CRC protected and self recovering GigaChip Interface
  - SEU resistant 1T-SRAM Memory core: < 10 FIT/Mb
  - Bit Safe™ Self Test and Self Repair Option

### Applications

<table>
<thead>
<tr>
<th>Applications</th>
<th>BE1 - MSR576</th>
<th>MSR620</th>
<th>MSR820</th>
<th>MSR720</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lookup, LPM, Hash</td>
<td>Highest single component Access Rate</td>
<td>Onboard ALU</td>
<td>Onboard ALU+</td>
<td></td>
</tr>
<tr>
<td>Statistics</td>
<td>Onboard ALU</td>
<td>Onboard ALU+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buffer – up to 80% efficient</td>
<td>Per cycle Burst, Write Broadcast</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metering, Dual Ops</td>
<td>Per cycle Burst, Write Broadcast</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semaphore, Link List</td>
<td>Fixed Macro</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State, Queuing, Link List</td>
<td>Atomic Ops</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dual Port w/ Data Coherency, 36 bit word access</td>
</tr>
</tbody>
</table>

Copyright ©MoSys, Inc. 2013. All rights reserved.
Functional Design
Bandwidth Engine MSR720 Architecture

Memory Operations:
- Rd/Wr 72b
- Wr 36b

Bank Conflict Resolution:
- 32K x 72b SRAM
- Delayed Write Cache

375 MHz Clock x
- 8 Reads +
- 8 Writes

→ 6GA Internally

4 x Partitions:
- 64 Single Port Banks
- 32K x 72b each

Copyright © MoSys, Inc. 2013. All rights reserved.
MSR720: Bandwidth Engine 2 – Access

- Simultaneous Read & Write @ same address
  - treat each partition as a single bank
  - >>2x the access performance of QDR SRAM
    - 4.5GA: 3 billion reads/sec, 1.5B 36b write/sec
    - 72b & 36b words each access
    - 12 ns read latency pin to pin
    - 2.7ns tRC cycle time
    - 8.5W @ 12.5G system power
  - 90% efficient transport protocol
    - Up to Sixteen 15G serial lanes (2 links of 8 lanes)

- High Density
  - 576Mbit 1T-SRAM ® memory core
    - 19mm x 19mm package – 1mm pitch

- High Reliability
  - 70X better SER than 6T-based SRAM
    - Memory core: < 10 FIT/Mb native
    - Interface: < 1 FIT
    - Bit Safe™ Self Test and Self Repair Option
MSR720: Basic ‘Dual Port’ Operation

- **Flat Partition Mode:**
  - Double the effective bandwidth in worst case:
    - Bank Conflict Resolution (BCR) function allows for simultaneous Read and Write of the same bank
    - Implements dual port behavior with single port banks
  - Guarantees data coherency
  - 3 billion accesses / sec @ 15G

---

**Memory Partition**

- banks hidden = ‘flat’ addressing
- BCR guarantees full data coherency

---

**Scheduling Balanced R:W**

<table>
<thead>
<tr>
<th>Frame</th>
<th>Input (RX)</th>
<th>Output (TX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0/1</td>
<td>QATX</td>
</tr>
<tr>
<td></td>
<td>RDFP</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>WRFP</td>
<td>RDFP</td>
</tr>
<tr>
<td>2</td>
<td>WD</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>2/3</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>RDFP</td>
<td>RDFP</td>
</tr>
<tr>
<td></td>
<td>WRFP</td>
<td>WRFP</td>
</tr>
<tr>
<td>4</td>
<td>WD</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>RDFP</td>
<td>RDFP</td>
</tr>
</tbody>
</table>

**Performance @ 15G**

- Write
- Read

---

Copyright ©MoSys, Inc. 2013. All rights reserved.
“BCR” Implemented w/Delayed Write Cache
Conceptual Block Diagram

Delayed Write Cache
Dual Port Memory

Write Address
Bank Offset

Addr Tag

Cached Write Data

WAB #
WAB Offset

RAB Offset

RAB #

Bank 0
Bank 1
Bank 2
Bank n-1
Bank n

Read Address
Bank Offset

Addr Tag

Cached Read Data

Hit

Read Address
Bank #

Read Data

Write Address
Bank #

Write Data

Read Data
The MSR720 supports Bank Aware commands.

- This can be used to improve the data read performance on the interface, however bypasses the BCR logic.
  
- Useful for unified memory applications combining dual port SRAM performance of “buffers” or “state” tables and read-only “lookup” tables.
  - The two address ranges cannot overlap.

MSR720 supports 36b write operations: WRFP and WRBA

- Useful for small word size tables (pointers)
- Increases write performance
Mapping Ports and Banks

For State Memory (WrFP, RdFP)

For other functions like Look Up Table (RdBA)
MSR720 Access Scheduling

- Flat Partition guarantees no bank conflict between read and write to any address, even in the same cycle

Balanced R:W Controller (Basic 72b R:W Operation)

<table>
<thead>
<tr>
<th>Frame</th>
<th>Input (RX)</th>
<th>Output (TX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMDARX</td>
<td>CMDBRX</td>
<td>QATX</td>
</tr>
<tr>
<td>1</td>
<td>RDFP WRFP RDFP WRFP</td>
<td>16 RDFP Data RDFP Data</td>
</tr>
<tr>
<td>2</td>
<td>WD WD</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>RDFP WRFP RDFP WRFP</td>
<td>18 RDFP Data RDFP Data</td>
</tr>
<tr>
<td>4</td>
<td>WD WD</td>
<td>19</td>
</tr>
</tbody>
</table>

Native BE Controller (72b WRITES)

<table>
<thead>
<tr>
<th>Frame</th>
<th>Input (RX)</th>
<th>Output (TX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMDARX</td>
<td>CMDBRX</td>
<td>QATX</td>
</tr>
<tr>
<td>1</td>
<td>RDFP RDBA RDFP RDBA</td>
<td>16 RDFP RDFP</td>
</tr>
<tr>
<td>2</td>
<td>WRFP WDL WRFP WDL</td>
<td>17 RDBA RDBA</td>
</tr>
<tr>
<td>3</td>
<td>RDFP RDBA RDFP RDBA</td>
<td>18 RDFP RDFP</td>
</tr>
<tr>
<td>4</td>
<td>WRFP WDL WRFP WDL</td>
<td>19 RDBA RDBA</td>
</tr>
</tbody>
</table>

Native BE Controller (36b WRITES)

<table>
<thead>
<tr>
<th>Frame</th>
<th>Input (RX)</th>
<th>Output (TX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMDARX</td>
<td>CMDBRX</td>
<td>QATX</td>
</tr>
<tr>
<td>1</td>
<td>RDFP RDBA</td>
<td>16 RDFP RDBA</td>
</tr>
<tr>
<td>2</td>
<td>WRFP WDL</td>
<td>17 RDBA RDBA</td>
</tr>
<tr>
<td>3</td>
<td>RDFP RDBA</td>
<td>18 RDFP RDBA</td>
</tr>
<tr>
<td>4</td>
<td>WRFP WDL</td>
<td>19 RDBA RDBA</td>
</tr>
</tbody>
</table>

Partition Access Restrictions

<table>
<thead>
<tr>
<th>Frame</th>
<th>RX</th>
</tr>
</thead>
<tbody>
<tr>
<td>0/1</td>
<td>P0 P1</td>
</tr>
<tr>
<td>2/3</td>
<td>P2 P3</td>
</tr>
</tbody>
</table>

Copyright ©MoSys, Inc. 2013. All rights reserved.
MSR 720: Breaking 4.5GA

- **Second Gen BE:** MSR 720 Peak w/15Gbps SerDes
- **First Gen BE**
- **RLDRAM3**
- **Sigma-Quad or QDR SRAM BL2 @ 600MHz**

*Note: Datasheet comparison - 1 access = 72 bits
- FP Access is arbitrated by BCR.
- BA Access is to/from array only. Must be bank aware.*
Get More Done With The Same Amount of I/O
Bandwidth Engine 2 Architecture & Family

- **Parallel Array Architecture** ... Performance up to:
  - 16 outstanding transactions
  - 4.5G Accesses per second
  - 192 Gbps full duplex throughput
  - ~12ns deterministic read latency
  - 2.7ns Random cycle time (tRC)

- **GigaChip Interface** ... 90% Efficient Transport Protocol
  - Up to sixteen low latency SerDes lanes (8G to 15G)

- **High Reliability** ... 70X better SER than 6T-SRAM
  - Full ECC support; 72bit array and macro datapath operations
  - CRC protected and self recovering GigaChip Interface
  - SEU resistant 1T-SRAM Memory core: < 10 FIT/Mb
  - Bit Safe™ Self Test and Self Repair Option

<table>
<thead>
<tr>
<th>Applications</th>
<th>BE1 - MSR576</th>
<th>MSR620</th>
<th>MSR820</th>
<th>MSR720</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lookup, LPM, Hash</td>
<td>Highest single component Access Rate</td>
<td>Onboard ALU</td>
<td>Onboard ALU+</td>
<td></td>
</tr>
<tr>
<td>Statistics</td>
<td>Onboard ALU</td>
<td>Per cycle Burst, Write Broadcast</td>
<td>Fixed Macro</td>
<td></td>
</tr>
<tr>
<td>Buffer – up to 80% efficient</td>
<td>Per cycle Burst, Write Broadcast</td>
<td></td>
<td>Fixed Macro</td>
<td>Atomic Ops</td>
</tr>
<tr>
<td>Metering, Dual Ops</td>
<td>Fixed Macro</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semaphore, Link List</td>
<td>Atomic Ops</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State, Queuing, Link List</td>
<td>Dual Port w/ Data Coherency, 36 bit word access</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Dual Counter: 5 Stage Pipeline

Includes Index Compare Logic (not illustrated) for Data Forwarding:

- In case of an index match, data is forwarded in the pipeline
- Prevents stale data in the pipeline
- Similar pipeline for Split Counter
- End-to-End ECC Protection
- 16 ns to completion
Individual Flow Programmability
- Meter Type
- Flow Rates
- Thresholds

8M Two Color Flows
4M Three Color Flows
Line Rate 4x100G

Basic Meter (Two Color)

Single Rate Three Color Meter

Two Rate Three Color Meter
Second Generation Bandwidth Engine Architecture

- Up to 4.5 billion external memory accesses per second with 16 SerDes Lanes
- Macros support up to 6 billion internal accesses per second with 8 SerDes Lane
- Macros execute Atomically: Stats, Metering, Read & Set, Test & Set

Supports:
- 4 x 100GE ports with Stats + Metering over 8 lanes

Intelligent Offload, Fire Forward Architecture
Physical Design
Conceptual Timing & Data Access Control

BE 1: Read Latency of 15.9ns vs. BE2: Read Latency of ~12.5ns

Data Path
- Rx & Bit Aligner
  - Analog
  - DeSerializer
  - DeScrambler
  - Slip Buf
  - Lane/FrameAligner
  - ALU
  - BCR
  - Memory Macro
  - Mem Access & RMW
  - Serializer & Tx

Timing Path
- Clock Data Recovery
- Core Clk
- Bit Clk
- Access Phase
- DLL Timing

BE 1: Read Latency of 15.9ns vs. BE2: Read Latency of ~12.5ns
MSR720 Layout

QDR like Dual Port Performance for 1.10x vs 3x die area cost

~6x Area

BCR Cache

32K x 72b

32K x (72b + 8b + 6b +1b)
Data + ECC + Tag + Valid

8x SerDes

SerDes I/O 5% of Die Area @ 240 Gbps
Thank you

Michael J. Miller

mmiller@mosys.com